> **This Raw Listing contains the General Information Section and up to first 5 pages.**

```
 1   <110> APPLICANT: Chen, Shiuan
 2         Zhou, Dujin
 3   <120> TITLE OF INVENTION: DRUG SCREENING USING A PROLINE-RICH NUCLEAR RECEPTOR
 4         CO-REGULATORY PROTEIN/NUCLEAR RECEPTOR CO-EXPRESSION
 5         SYSTEM
 6   <130> FILE REFERENCE: 2124-311
 7   <140> CURRENT APPLICATION NUMBER: US/09/552,705
 8   <141> CURRENT FILING DATE: 2000-04-19
 9   <150> EARLIER APPLICATION NUMBER: U.S. 60/129,873
10   <151> EARLIER FILING DATE: 1999-04-19
11   <160> NUMBER OF SEQ ID NOS: 9
12   <170> SOFTWARE: PatentIn Ver. 2.0
13   <210> SEQ ID NO 1
14   <211> LENGTH: 30
15   <212> TYPE: DNA
16   <213> ORGANISM: Artificial Sequence
17   <220> FEATURE:
18   <223> OTHER INFORMATION: Description of Artificial Sequence:Sense primer.
19   <400> SEQUENCE: 1
20         gccgaattcg gggagggcag gggtgaagtg                              30
21   <210> SEQ ID NO 2
22   <211> LENGTH: 39
23   <212> TYPE: DNA
24   <213> ORGANISM: Artificial Sequence
25   <220> FEATURE:
26   <223> OTHER INFORMATION: Description of Artificial Sequence:Antisense
27         primer.
28   <400> SEQUENCE: 2
29         ggcgtcgacg gatcctcaga ctgtggcagg gaaaccctc                    39
30   <210> SEQ ID NO 3
31   <211> LENGTH: 12
32   <212> TYPE: DNA
33   <213> ORGANISM: Homo sapiens
34   <400> SEQUENCE: 3
35         ccaaggtcag aa                                                 12
36   <210> SEQ ID NO 4
37   <211> LENGTH: 5
38   <212> TYPE: PRT
39   <213> ORGANISM: Homo sapiens
40   <400> SEQUENCE: 4
41         Leu Lys Thr Leu Leu
42           1               5
43   <210> SEQ ID NO 5
44   <211> LENGTH: 7
```

PAGE: 2   **RAW SEQUENCE LISTING**   DATE: 05/08/2000
PATENT APPLICATION **US/09/552,705** TIME: 18:55:23

Input Set: **I552705.RAW**

```
45    <212> TYPE: PRT
46    <213> ORGANISM: Homo sapiens
47    <400> SEQUENCE: 5
48         Ser Asp Pro Pro Ser Pro Ser
49           1               5
50    <210> SEQ ID NO 6
51    <211> LENGTH: 7
52    <212> TYPE: PRT
53    <213> ORGANISM: Homo sapiens
54    <220> FEATURE:
55    <221> NAME/KEY: MUTAGEN
56    <222> LOCATION: (3)
57    <223> OTHER INFORMATION: This residue has been changed from a proline.
58    <220> FEATURE:
59    <221> NAME/KEY: MUTAGEN
60    <222> LOCATION: (6)
61    <223> OTHER INFORMATION: This residue has been changed from a proline.
62    <400> SEQUENCE: 6
63         Ser Asp Ala Pro Ser Ala Ser
64           1               5
65    <210> SEQ ID NO 7
66    <211> LENGTH: 2061
67    <212> TYPE: DNA
68    <213> ORGANISM: Homo sapiens
69    <220> FEATURE:
70    <221> NAME/KEY: CDS
71    <222> LOCATION: (114)..(1094)
72    <400> SEQUENCE: 7
73         tgttccgcga tcttctcagg ctctcctagc agcatccatc gccgccaccc tatcttcact 60
74         ggcttcacct tctccttctc tcttcgttgc tgagcgacaa gcttcctagc gct atg      116
75                                                                         Met
76                                                                          1
77         act gtc gtc tcc gtc ccg cag cgg gag ccg ctc gtc ctg ggt ggc cgc   164
78         Thr Val Val Ser Val Pro Gln Arg Glu Pro Leu Val Leu Gly Gly Arg
79                           5               10                  15
80         ctt gcg ccg ctt ggc ttt tcc tcc cga ggt tac ttt ggg gcc ctc ccg   212
81         Leu Ala Pro Leu Gly Phe Ser Ser Arg Gly Tyr Phe Gly Ala Leu Pro
82                      20              25              30
83         atg gtg acc acg gct ccg cct cct tta ccc cgg atc ccg gac ccc cgg   260
84         Met Val Thr Thr Ala Pro Pro Pro Leu Pro Arg Ile Pro Asp Pro Arg
85                  35              40                  45
86         gca ctg ccc ccg acc ctc ttc ctc cct cat ttc cta ggg gga gat ggc   308
87         Ala Leu Pro Pro Thr Leu Phe Leu Pro His Phe Leu Gly Gly Asp Gly
88            50              55                  60                  65
89         ccg tgt ctg acc ccc cag cct cgc gct cca gca gct ctg ccc aac cgc   356
90         Pro Cys Leu Thr Pro Gln Pro Arg Ala Pro Ala Ala Leu Pro Asn Arg
91                      70              75                  80
92         agc ctc gcc gtg gcg gga ggc act cct cgg gca gcg ccg aag aag cgg   404
93         Ser Leu Ala Val Ala Gly Gly Thr Pro Arg Ala Ala Pro Lys Lys Arg
94                  85              90                  95
```

```
 95        cga aag aag aag gtg cgg gcc agc ccc gca ggg cag ctg ccc agc cgc   452
 96        Arg Lys Lys Lys Val Arg Ala Ser Pro Ala Gly Gln Leu Pro Ser Arg
 97                100                 105                 110
 98        ttc cac cag tac cag cag cac cgg ccg agt ctg gag ggc ggc cgg agc   500
 99        Phe His Gln Tyr Gln Gln His Arg Pro Ser Leu Glu Gly Gly Arg Ser
100                115                 120                 125
101        ccc gcg acc ggc ccg agc gga gcg cag gag gtc ccg ggc ccg gcc gcc   548
102        Pro Ala Thr Gly Pro Ser Gly Ala Gln Glu Val Pro Gly Pro Ala Ala
103                130                 135                 140                 145
104        gcc ttg gcc ccg agt cct gca gcc gca gcc ggc acg gag gga gcc agc   596
105        Ala Leu Ala Pro Ser Pro Ala Ala Ala Ala Gly Thr Glu Gly Ala Ser
106                            150                 155                 160
107        ccc gac ctt gcc ccg ctg cgg ccc gcg gct ccc ggc caa acc ccc ctc   644
108        Pro Asp Leu Ala Pro Leu Arg Pro Ala Ala Pro Gly Gln Thr Pro Leu
109                165                 170                 175
110        agg aaa gag gtt tta aaa tca aag atg gga aaa tcg gag aaa att gcc   692
111        Arg Lys Glu Val Leu Lys Ser Lys Met Gly Lys Ser Glu Lys Ile Ala
112                180                 185                 190
113        ctt ccc cat ggc cag ctt gtt cat ggt ata cac ttg tat gag caa cca   740
114        Leu Pro His Gly Gln Leu Val His Gly Ile His Leu Tyr Glu Gln Pro
115                195                 200                 205
116        aag ata aac aga cag aaa agc aaa tat aac ttg cca cta acc aag atc   788
117        Lys Ile Asn Arg Gln Lys Ser Lys Tyr Asn Leu Pro Leu Thr Lys Ile
118                210                 215                 220                 225
119        acc tct gca aaa aga aat gaa aac aac ttt tgg cag gat tct gtt tca   836
120        Thr Ser Ala Lys Arg Asn Glu Asn Asn Phe Trp Gln Asp Ser Val Ser
121                            230                 235                 240
122        tct gac aga att cag aag cag gaa aaa aag cct ttt aaa aat acc gag   884
123        Ser Asp Arg Ile Gln Lys Gln Glu Lys Lys Pro Phe Lys Asn Thr Glu
124                245                 250                 255
125        aac att aaa aat tcg cat ttg aag aaa tca gca ttt cta act gaa gtg   932
126        Asn Ile Lys Asn Ser His Leu Lys Lys Ser Ala Phe Leu Thr Glu Val
127                260                 265                 270
128        agc caa aag gaa aat tat gct ggg gca aag ttt agt gat cca cct tct   980
129        Ser Gln Lys Glu Asn Tyr Ala Gly Ala Lys Phe Ser Asp Pro Pro Ser
130                275                 280                 285
131        cct agt gtt ctt cca aag cct cct agt cac tgg atg gga agc act gtt   1028
132        Pro Ser Val Leu Pro Lys Pro Pro Ser His Trp Met Gly Ser Thr Val
133                290                 295                 300                 305
134        gaa aat tcc aac caa aac agg gag ctg atg gca gta cac tta aaa acc   1076
135        Glu Asn Ser Asn Gln Asn Arg Glu Leu Met Ala Val His Leu Lys Thr
136                            310                 315                 320
137        ctc ctc aaa gtt caa act tagatttcag atttcagtat gtgtgtaaaa   1124
138        Leu Leu Lys Val Gln Thr
139                325
140        cataattttt cccatatccc tggactcttg agaaaattgg tacagaaatg gaaatttgcc 1184
141        ttgttgcaac atacaattgc aaaagatgag tttaaaaaat tacatacaaa cagcttgtat 1244
142        tatattttat attttgtaaa tactgtatac catgtattat gtgtatattg ttcatacttg 1304
143        agaggtatat tatagttttg ttatgaaagt atgtattttg ccctgcccac attgcaggtg 1364
144        ttttgtatat atacaatgga taaattttaa gtgtgtgcta aggcacatgg aagaccgatt 1424
```

```
145        ttatttgcac aaggtactga gattttttc aagaaacagc tgtcaaatct caaggtgaag 1484
146        atctaaatgt gaacagttta ctaatgcact actgaagttt aaatctgtgg cacaatcaat 1544
147        gtaagcatgg ggtttgtttc tctaaattga tttgtaatct gaaattactg aacaactcct 1604
148        attcccattt ttgctaaact caatttctgg ttttggtata tatccattcc agcttaatgc 1664
149        ctctaatttt aatgccaaca aaattggttg taatcaaatt ttaaaataat aataatttgg 1724
150        cccccccttt taaaatagtc ttgactcttt gtgtgtgact gtttctcatg tttgaatgtg 1784
151        tgactaggag atgattttgt gtggttggat tttttttgact tctactttac tggctgagtg 1844
152        tgagccgcca tgcctggcca taatctacat tttcttacca ggagcagcat tgaggttttt 1904
153        gagcatagta cttgactact ctagaggctg agacgggagc atctcttgag cctgagaagt 1964
154        ggagattgca attgagctag gatcaggcca ctgcactcca gcctgggtaa cagacgctgt 2024
155        ctcaaaaaaa aggccaagag aaagtaaggg agacaga                            2061
156   <210> SEQ ID NO 8
157   <211> LENGTH: 327
158   <212> TYPE: PRT
159   <213> ORGANISM: Homo sapiens
160   <400> SEQUENCE: 8
161        Met Thr Val Val Ser Val Pro Gln Arg Glu Pro Leu Val Leu Gly Gly
162         1               5                  10                  15
163        Arg Leu Ala Pro Leu Gly Phe Ser Ser Arg Gly Tyr Phe Gly Ala Leu
164                       20                  25                  30
165        Pro Met Val Thr Thr Ala Pro Pro Pro Leu Pro Arg Ile Pro Asp Pro
166                       35                  40                  45
167        Arg Ala Leu Pro Pro Thr Leu Phe Leu Pro His Phe Leu Gly Gly Asp
168              50                  55                  60
169        Gly Pro Cys Leu Thr Pro Gln Pro Arg Ala Pro Ala Ala Leu Pro Asn
170           65                  70                  75                  80
171        Arg Ser Leu Ala Val Ala Gly Gly Thr Pro Arg Ala Ala Pro Lys Lys
172                       85                  90                  95
173        Arg Arg Lys Lys Lys Val Arg Ala Ser Pro Ala Gly Gln Leu Pro Ser
174                      100                 105                 110
175        Arg Phe His Gln Tyr Gln Gln His Arg Pro Ser Leu Glu Gly Gly Arg
176                      115                 120                 125
177        Ser Pro Ala Thr Gly Pro Ser Gly Ala Gln Glu Val Pro Gly Pro Ala
178              130                 135                 140
179        Ala Ala Leu Ala Pro Ser Pro Ala Ala Ala Ala Gly Thr Glu Gly Ala
180           145                 150                 155                 160
181        Ser Pro Asp Leu Ala Pro Leu Arg Pro Ala Ala Pro Gly Gln Thr Pro
182                      165                 170                 175
183        Leu Arg Lys Glu Val Leu Lys Ser Lys Met Gly Lys Ser Glu Lys Ile
184                      180                 185                 190
185        Ala Leu Pro His Gly Gln Leu Val His Gly Ile His Leu Tyr Glu Gln
186                      195                 200                 205
187        Pro Lys Ile Asn Arg Gln Lys Ser Lys Tyr Asn Leu Pro Leu Thr Lys
188              210                 215                 220
189        Ile Thr Ser Ala Lys Arg Asn Glu Asn Asn Phe Trp Gln Asp Ser Val
190        225                 230                 235                 240
191        Ser Ser Asp Arg Ile Gln Lys Gln Glu Lys Lys Pro Phe Lys Asn Thr
192                      245                 250                 255
193        Glu Asn Ile Lys Asn Ser His Leu Lys Lys Ser Ala Phe Leu Thr Glu
194                      260                 265                 270
```

PAGE: 5

RAW SEQUENCE LISTING
PATENT APPLICATION **US/09/552,705**

DATE: 05/08/2000
TIME: 18:55:23

Input Set: **I552705.RAW**

```
195          Val Ser Gln Lys Glu Asn Tyr Ala Gly Ala Lys Phe Ser Asp Pro Pro
196                  275                 280                 285
197          Ser Pro Ser Val Leu Pro Lys Pro Pro Ser His Trp Met Gly Ser Thr
198                  290                 295                 300
199          Val Glu Asn Ser Asn Gln Asn Arg Glu Leu Met Ala Val His Leu Lys
200          305                 310                 315                 320
201          Thr Leu Leu Lys Val Gln Thr
202                          325
203    <210> SEQ ID NO 9
204    <211> LENGTH: 23
205    <212> TYPE: PRT
206    <213> ORGANISM: Homo sapiens
207    <400> SEQUENCE: 9
208          Asn Tyr Ala Gly Ala Lys Phe Ser Asp Pro Pro Ser Pro Ser Val Leu
209            1           5                 10                  15
210          Pro Lys Pro Pro Ser His Trp
211                          20
```

Line ? Error/Warning   Original Text
----- - -- --